

SEQUENCE LISTING

SEQ ID NO:1
SEQUENCE LENGTH: 5399
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE
ORGANISM: human
CELL LINE: KG-1

A GAA TTT TCA AAT CCT GAA ACT CAG AAT CTG GAT GCC ATG GAA CAA GTT 49
Glu Phe Ser Asn Pro Glu Thr Gln Asn Leu Asp Ala Met Glu Gln Val
1 5 10 15
GGT CTG GAA TCC TTA CAG TTT GAC TAT CCT GGT AAT CAG GTA CCA ATG 97
Gly Leu Glu Ser Leu Gln Phe Asp Tyr Pro Gly Asn Gln Val Pro Met
20 25 30
GAC TCT TCA GGA GCT ACT GTA GGC CTT TTT GAC TAC AAT TCC CAG CAG 145
Asp Ser Ser Gly Ala Thr Val Gly Leu Phe Asp Tyr Asn Ser Gln Gln
35 40 45
CAG CTC TTT CAG AGG ACT AAT GCA CTA ACA GTT CAA CAG TTA ACT GCA 193
Gln Leu Phe Gln Arg Thr Asn Ala Leu Thr Val Gln Gln Leu Thr Ala
50 55 60
GCT CAA CAG CAG CAA TAT GCA TTA GCA GCA GCT CAG CAG CCA CAT ATA 241
Ala Gln Gln Gln Gln Tyr Ala Leu Ala Ala Ala Gln Gln Pro His Ile
65 70 75 80
GCT GGT GTA TTC TCA GCA GGC CTT GCT CCA GCT GCA TTT GTG CCA AAT 289
Ala Gly Val Phe Ser Ala Gly Leu Ala Pro Ala Ala Phe Val Pro Asn
85 90 95
CCA TAC ATT ATT AGT GCT GCT CCT CCA GGG ACC GAT CCG TAT ACT GCA 337
Pro Tyr Ile Ile Ser Ala Ala Pro Pro Gly Thr Asp Pro Tyr Thr Ala
100 105 110
GCA GGA TTG GCT GCA GCA GCT ACA TTA GCA GGT CCA GCA GTG GTT CCA 385
Ala Gly Leu Ala Ala Ala Ala Thr Leu Ala Gly Pro Ala Val Val Pro
115 120 125
CCT CAG TAT TAC GGC GTT CCA TGG GGG GTG TAT CCA GCC AAC TTA TTT 433
Pro Gln Tyr Tyr Gly Val Pro Trp Gly Val Tyr Pro Ala Asn Leu Phe
130 135 140
CAG CAG CAA GCT GCA GCT GCG GCA AAT AAC ACA GCC AGT CAG CAA GCA 481
Gln Gln Gln Ala Ala Ala Ala Asn Asn Thr Ala Ser Gln Gln Ala
145 150 155 160
GCA TCA CAA GCT CAG CCT GGA CAG CAA CAG GTT CTC CGT GCT GGA GCA 529
Ala Ser Gln Ala Gln Pro Gly Gln Gln Gln Val Leu Arg Ala Gly Ala
165 170 175
GGT CAG CGT CCT CTT ACT CCC AAT CAG GGT CAG CAA GGG CAG CAA GCA 577
Gly Gln Arg Pro Leu Thr Pro Asn Gln Gly Gln Gln Gly Gln Ala
180 185 190
GAA TCA CTT GCG GCA GCT GCA GCA GCA AAT CCA ACA TTG GCT TTT GGT 625
Glu Ser Leu Ala Ala Ala Ala Ala Ala Asn Pro Thr Leu Ala Phe Gly
195 200 205
CAG GGT CTT GCT ACT GGC ATG CCA GGC TAT CAA GTA CTA GCT CCA ACT 673

Gln	Gly	Leu	Ala	Thr	Gly	Met	Pro	Gly	Tyr	Gln	Val	Leu	Ala	Pro	Thr	
210						215					220					
GCC	TAT	TAT	GAT	CAG	ACT	GGT	GCC	TTA	GTG	GTT	GGC	CCT	GGA	GCA	AGG	721
Ala	Tyr	Tyr	Asp	Gln	Thr	Gly	Ala	Leu	Val	Val	Gly	Pro	Gly	Ala	Arg	
225					230						235				240	
ACT	GGC	CTT	GGA	GCT	CCA	GTT	CGG	TTA	ATG	GCT	CCA	ACA	CCT	GTT	TTA	769
Thr	Gly	Leu	Gly	Ala	Pro	Val	Arg	Leu	Met	Ala	Pro	Thr	Pro	Val	Leu	
				245					250						255	
ATT	AGT	TCA	GCA	GCA	GCA	CAA	GCT	GCA	GCA	GCA	GCA	GCA	GCT	GGA	GGA	817
Ile	Ser	Ser	Ala	Ala	Ala	Gln	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Gly	
			260						265						270	
ACT	GCA	AGT	AGC	CTT	ACA	GGC	AGC	ACA	AAT	GGT	CTG	TTT	CGG	CCA	ATT	865
Thr	Ala	Ser	Ser	Leu	Thr	Gly	Ser	Thr	Asn	Gly	Leu	Phe	Arg	Pro	Ile	
			275				280								285	
GGC	ACT	CAG	CCA	CCA	CAG	CAG	CAA	CAG	CAG	CCA	AGC	ACT	AAT	CTG		913
Gly	Thr	Gln	Pro	Pro	Gln	Gln	Gln	Gln	Gln	Pro	Ser	Thr	Asn	Leu		
			290				295								300	
CAA	TCT	AAT	TCA	TTT	TAT	GGA	AGC	AGT	TGT	TTG	ACT	AAT	AGC	TCC	CAG	961
Gln	Ser	Asn	Ser	Phe	Tyr	Gly	Ser	Ser	Ser	Leu	Thr	Asn	Ser	Ser	Gln	
305					310										320	
AGT	AGT	TCT	TTA	TTT	TCT	CAT	GGA	CCT	GGT	CAA	CCT	GGA	AGT	ACA	TCT	1009
Ser	Ser	Ser	Leu	Phe	Ser	His	Gly	Pro	Gly	Gln	Pro	Gly	Ser	Thr	Ser	
				325					330						335	
CTT	GGC	TTT	GGA	AGT	GGT	AAC	TCT	TTG	GGT	GCT	GCT	ATA	GGC	TCA	GCC	1057
Leu	Gly	Phe	Gly	Ser	Gly	Asn	Ser	Leu	Gly	Ala	Ala	Ile	Gly	Ser	Ala	
			340						345						350	
CTC	AGT	GGA	TTT	GGT	TCA	TCA	GTT	GGC	AGT	TCT	GCA	AGT	AGT	AGT	GCC	1105
Leu	Ser	Gly	Phe	Gly	Ser	Ser	Val	Gly	Ser	Ser	Ala	Ser	Ser	Ser	Ala	
			355					360							365	
ACA	AGG	AGA	GAG	TCT	CTA	TCT	ACT	AGC	TCT	GAC	TTG	TAC	AAA	AGA	TCT	1153
Thr	Arg	Arg	Glu	Ser	Leu	Ser	Thr	Ser	Ser	Asp	Leu	Tyr	Lys	Arg	Ser	
						375					380					
AGT	AGC	AGC	CTA	GCA	CCC	ATA	GGG	CAA	CCA	TTT	TAC	AAT	AGT	CTG	GGA	1201
Ser	Ser	Ser	Leu	Ala	Pro	Ile	Gly	Gln	Pro	Phe	Tyr	Asn	Ser	Leu	Gly	
385					390										400	
TTT	TCC	TCC	TCT	CCA	AGT	CCA	ATA	GGC	ATG	CCT	CTG	CCA	AGC	CAA	ACT	1249
Phe	Ser	Ser	Ser	Pro	Ser	Pro	Ile	Gly	Met	Pro	Leu	Pro	Ser	Gln	Thr	
				405					410						415	
CCA	GGA	CAT	TCA	CTT	ACG	CCA	CGG	CCA	TCA							

Leu	Arg	Tyr	Asn	Arg	Ser	Asp	Ile	Met	Pro	Ser	Gly	Arg	Ser	Arg	Leu	
TTG	GAA	GAT	TTC	AGA	AAC	AAC	CGC	TTC	CCA	AAC	CTT	CAG	CTT	AGA	GAC	1537
Leu	Glu	Asp	Phe	Arg	Asn	Asn	Arg	Phe	Pro	Asn	Leu	Gln	Leu	Arg	Asp	
TTG	ATT	GGA	CAT	ATA	GTT	GAG	TTT	TCT	CAA	GAC	CAG	CAT	GGT	TCT	AGA	1585
Leu	Ile	Gly	His	Ile	Val	Glu	Phe	Ser	Gln	Asp	Gln	His	Gly	Ser	Arg	
TTC	ATA	CAG	CAA	AAA	CTA	GAG	AGA	GCT	AGT	CCA	GCT	GAG	CGA	CAG	ATG	1633
Phe	Ile	Gln	Gln	Lys	Leu	Glu	Arg	Ala	Thr	Pro	Ala	Glu	Arg	Gln	Met	
GTA	TTT	AAT	GAA	ATT	CTG	CAA	GCA	GCC	TAT	CAA	TTA	ATG	ACT	GAT	GTT	1681
Val	Phe	Asn	Glu	Ile	Leu	Gln	Ala	Ala	Tyr	Gln	Leu	Met	Thr	Asp	Val	
TTT	GGC	AAC	TAT	GTT	ATA	CAG	AAG	TTT	TTT	GAG	TTT	GGG	AGT	CTG	GAT	1729
Phe	Gly	Asn	Tyr	Val	Ile	Gln	Lys	Phe	Phe	Glu	Phe	Gly	Ser	Leu	Asp	
CAA	AAA	TTA	GCC	CTG	GCT	ACT	CGT	ATT	CGT	GGT	CAT	GTT	CTA	CCC	TTA	1777
Gln	Lys	Leu	Ala	Leu	Ala	Thr	Arg	Ile	Arg	Gly	His	Val	Leu	Pro	Leu	
GCC	TTG	CAG	ATG	TAT	GGC	TGC	CGC	GTT	ATT	CAG	AAA	GCA	TTA	GAA	TCT	1825
Ala	Leu	Gln	Met	Tyr	Gly	Cys	Arg	Val	Ile	Gln	Lys	Ala	Leu	Glu	Ser	
ATT	TCT	TCT	GAC	CAG	CAG	AGT	GAA	ATG	GTA	AAG	GAG	CTG	GAT	GGT	CAT	1873
Ile	Ser	Ser	Asp	Gln	Gln	Ser	Glu	Met	Val	Lys	Glu	Leu	Asp	Gly	His	
GTG	CTC	AAA	TGT	GTG	AAA	GAT	CAG	AAT	GGA	AAC	CAT	GTT	GTA	CAA	AAA	1921
Val	Leu	Lys	Cys	Val	Lys	Asp	Gln	Asn	Gly	Asn	His	Val	Val	Gln	Lys	
TGT	ATC	GAA	TGT	GTT	CAG	CCA	CAG	TCA	CTA	CAG	TTC	ATC	ATT	GAT	GCT	1969
Cys	Ile	Glu	Cys	Val	Gln	Pro	Gln	Ser	Leu	Gln	Phe	Ile	Ile	Asp	Ala	
TTC	AAG	GGA	CAA	GTA	TTT	GTG	CTT	TCA	ACT	CAT	CCT	TAT	GGC	TGC	AGA	2017
Phe	Lys	Gly	Gln	Val	Phe	Val	Leu	Ser	Thr	His	Pro	Tyr	Gly	Cys	Arg	
GTA	ATT	CAG	CGC	ATC	CTA	GAG	GAT	TGC	ACT	GCA	GAA	CAG	ACC	TTA	CCT	2065
Val	Ile	Gln	Arg	Ile	Leu	Glu	His	Cys	Thr	Ala	Glu	Gln	Thr	Leu	Pro	
ATC	TTA	GAA	GAA	CTC	CAC	CAA	CAT	ACA	GAG	CAG	TTG	GTA	CAG	GAT	CAG	2113
Ile	Leu	Glu	Glu	Leu	His	Gln	His	Thr	Glu	Gln	Leu	Val	Gln	Asp	Gln	
TAT	GGC	AAT	TAT	GTT	ATT	CAG	CAT	GTA	CTG	GAA	CAC	GGT	CGA	CCT	GAA	2161
Tyr	Gly	Asn	Tyr	Val	Ile	Gln	His	Val	Leu	Glu	His	Gly	Arg	Pro	Glu	
GAC	AAG	AGC	AAA	ATT	GTT	TCC	GAA	ATC	AGG	GGA	AAG	GTT	TTA	GCC	CTG	2209
Asp	Lys	Ser	Lys	Ile	Val	Ser	Glu	Ile	Arg	Gly	Lys	Val	Leu	Ala	Leu	
AGT	CAA	CAC	AAA	TTT	GCC	AGC	AAT	GTA	GTA	GAA	AAG	TGT	GTT	ACT	CAT	2257
Ser	Gln	His	Lys	Phe	Ala	Ser	Asn	Val	Val	Glu	Lys	Cys	Val	Thr	His	
GCC	TCC	CGT	GCT	GAG	AGA	GCT	TTA	CTG	ATT	GAC	GAG	GTT	TGC	TGC	CAG	2305

Ala Ser Arg Ala Glu Arg Ala Leu Leu Ile Asp Glu Val Cys Cys Gln
755 760 765
AAT GAT GGT CCT CAC AGT GCC TTA TAC ACC ATG ATG AAG GAC CAG TAT 2353
Asn Asp Gly Pro His Ser Ala Leu Tyr Thr Met Met Lys Asp Gln Tyr
770 775 780
GCC AAT TAC GTG GTT CAA AAG ATG ATT GAT ATG GCT GAA CCT GCT CAG 2401
Ala Asn Tyr Val Val Gln Lys Met Ile Asp Met Ala Glu Pro Ala Gln
785 790 795 800
AGA AAG ATA ATC ATG CAC AAG ATT CGA CCT CAC ATT ACT ACT TTG CGC 2449
Arg Lys Ile Ile Met His Lys Ile Arg Pro His Ile Thr Thr Leu Arg
805 810 815
AAA TAC ACA TAC GGG AAG CAT ATA CTG GCC AAG TTG GAA AAG TAT TAT 2497
Lys Tyr Thr Tyr Gly Lys His Ile Leu Ala Lys Leu Glu Lys Tyr Tyr
820 825 830
TTG AAG AAT AGC CCG GAC CTA GGA CCT ATT GGA GGA CCA CCA AAT GGA 2545
Leu Lys Asn Ser Pro Asp Leu Gly Pro Ile Gly Gly Pro Pro Asn Gly
835 840 845
ATG CTG TAAATTACAG GAGCAAGAGA AAGAAGATAA TTTAACCATG TGAAAAGAAT 2601
Met Leu 850

TTTTTTGTGC GTGAATTATC AAAACACAAC TCAACTATGA ATCTTCAATT TTTTTTTAAA 2661
GCAAAACTAT TTATTGACTT TATTCATCCA TTTGTAAATT TTTTAAGGTT CTTGTGTATA 2771
TTTGGGGGGT GGGGGATGAA TTATAAATTA TATTCAGCCC TGAGTGGAGA CCTATCAGAT 2781
TGGATTGCTG GCAAAGCACA GAATGCCTGT ATATGATGTA ACTGTATCAA AAATAAAAAG 2841
CTGTACATA TTTTGTAAT TTTTACCTTG TAAAGTCACA AAAATAGTTT TTAAAGGAAA 2901
AAGTACAGTA TTCTTTTAAT AAACCTGGCTC ACAGTCTGGT AGGTCTACAA CCCCATAGCA 2961
CAACAGGTTT ATAGAGATGT ATATAGAATT ATAGTCTTAA TTTTCTTCTT TTGCGTGAAA 3021
CCTTTTATAA CAGATTAACA ATCAACTGCA TAAATATTAT TAATATTTTA AAAAGAGTTA 3081
AGTTGTATTT TGATAATTCA CAAACTATCA TGCAATAAAC GAGTAAGTAG ACAAGAATAA 3141
AGTGGTTTGA GATGAAAAGA ACCTAACATT ATTTACAGTA GATGTGGTTT TAATACAATT 3201
ACTGCCCTAA AATGTCTCTG GCAATGTACA GAAATATTGT ATATACTTAC ATATGTAATT 3261
GTTGTAAGAG TTAAATACAA AATCATGGTG ACACTTCCAA TTAAGTGCAC TAAATGAAAA 3321
GTAAAGTCAC TTATTAACCT TTCAGTTTGG TTTGCAATGA GAAAGAGTGG AAATTTGTAT 3381
TTTGTTTTGC TTATAGAATT ACAGACATGT TGAGGAAGTG TTGAGCTTTA TTTTGCTTTT 3441
TCATAGAGGC AGAAAGTAGG AACCAGATAG AGATGAAAAG GGGCCACTGA AAAGTGAATT 3501
TGATAGCTCA GCATTTAAGC ATGATTACAT ATTCAGATAG CTCTTTTTCG TTTCTATAAA 3561
TATATGCATT GTGTGTGTAG TAATAGATGT AAGTTTACAC TTTGAAAGGA AATCTTGTTT 3621
CAATGTTTAT TATAAAAGCC TTGCTAATTT AGTAGTGATG CTTTCCTTGG TTGTACAGGT 3681
GTACATTTGT AAACCTTCAT GCTGTAAATG GAATTTGTTT TATCTCTTTG GGATACATTT 3741
GCATTTTAGT GTACATTTAC GTCCCTGCCC TCTTTGACCT GGCAATATAG TGTGTATAA 3801
TGTAATTTA TTTCTCCAAA TCGAGAGTGA TTTTAAAAA ATTTTATATC TTTATATGGT 3861
TTCAGAAGTA TGAACAGCT TTCTTTTTAT TATTGTGAGA TCATTTTGTT TTATAACATA 3921
GTTGTTGACT GTTAATATGG ACCTGCTAGA ATTTGGATCA CTTTCAATTG AAGTCAGGGT 3981
ATTGTGCATA ATAGAAAGTA TTGGACTGAG ATATTTGGTT ACCATGGAGG CCAATGCTTT 4041
TTTCATCTTA TTAAATGTGA TGTGACTTTT TTCTTTGTAC AGAAGAGTAC TGTATTTTGT 4101
AATAGCCTAC TCCCAAGTAA GAGCAAATCT GTATGATAAC ATTTTTCCT CTGGACATAA 4161
GACATAACAG TAACACGATG TACATTTAGA AGCGGCCTTA TGTACATTTT CCAACAATCT 4221
TTTTAAGGCA AAATTGTGAC CATATGTGTA TAATTAATAA CGTTTTTAAT CCTTGCCTA 4281
TGAAAATATT TTGGAAAAAA ACTTGCTGTG TATATTCAGT TTCTGAAAGA TAAAGAAAGT 4341
GCTTTGTATT TTGTTGAAGT CAGTATTTTG TATAAACATT TATGTTGACC CACTTATGTT 4401
CAGTGCTGAA AACTAAAATG AACATGCTAT TCTGTCAGCT GAATATGGAA GAGATCTTTT 4461

TTTACTAGAG ATCTGCAGAA GAAACGCAAT CTTCTGAGCA CAATATGGAA TCTAAAGGTT 4521
 TTATCACTTA GTTGTTTATA TTATGAACCT AAAAATAATG GCATAAAGTT TGGGGATGCC 4581
 AGGCATACTT TTTTCATGTTT GGTGTTGAGT TATTTTACTT TTCTAACCCA ACATTCCTTG 4641
 GTGAGACCAT TAAATCCAAA CACTTGTCAC CGTTCCTTCT CATAGTCACT CTGGGTCATC 4701
 AGCATGTCCC AGTCACTGCA GCAACGCCTT GTGTTTGTTT CATTTTTTTT AAACCCACAC 4761
 AAAGCCGCTG TCTCACTTTT TCCTACTTTA CCAACCTCAG AGTATTTTCGG CCCGTATCGA 4821
 ACTTTTGTTT TCAGTATCAG CCCATGGTTT CAGGATCAAA GCTGTCATGT TGGAGATTGG 4881
 TAATGGCTTT CCTGTCTTTG TACAGTTGAA TTCCTAGTCT TGCTTCATCC TTGCCCTCTG 4941
 TTGGCACAGG CATTATCTCT GCAATTTTATG AAAATGACAA GTAGAGAATA CTACATTGAG 5001
 AAATAAACC CTCTTCTTGG GGTCCTGATA CTCATTCCCA TTTGTCCCAG TGCTGACAAC 5061
 CCAATCTTCC CAATACTTTC AGGCCTGCTC TACAAAAGTA CCTGTTCTTG TAGAAATTTT 5121
 ACAGTCTGCC ATTTTGGGTG CCCACCCCAA TTTTACCTT TTAGTAAGTT GGCATGAAAT 5181
 TTTGGTAAAA TCTGAAAATC ACATTTTACA ATAAACAAT TGGGCAAAAC TACCTAGGCT 5241
 TTACTCTTGA GTGTCTCCTT TTGATAGGGA TTGTTTCTGG ACCAGTTTGT CTAAGTCCTG 5301
 GCTCTTATTG GTTCATATGA AATAATGTGA ACTTCACCTC TTTGTATATT ATGTATAAAT 5361
 TAGAAAATGA AAAATGTGTG AATAACATTG TATGAAAT 5399

SEQ ID NO:2
 SEQUENCE LENGTH: 850
 SEQUENCE TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE
 ORGANISM: human
 CELL LINE: KG-1

Glu Phe Ser Asn Pro Glu Thr Gln Asn Leu Asp Ala Met Glu Gln Val
 1 5 10 15
 Gly Leu Glu Ser Leu Gln Phe Asp Tyr Pro Gly Asn Gln Val Pro Met
 20 25 30
 Asp Ser Ser Gly Ala Thr Val Gly Leu Phe Asp Tyr Asn Ser Gln Gln
 35 40 45
 Gln Leu Phe Gln Arg Thr Asn Ala Leu Thr Val Gln Gln Leu Thr Ala
 50 55 60
 Ala Gln Gln Gln Gln Tyr Ala Leu Ala Ala Ala Gln Gln Pro His Ile
 65 70 75 80
 Ala Gly Val Phe Ser Ala Gly Leu Ala Pro Ala Ala Phe Val Pro Asn
 85 90 95
 Pro Tyr Ile Ile Ser Ala Ala Pro Pro Gly Thr Asp Pro Tyr Thr Ala
 100 105 110
 Ala Gly Leu Ala Ala Ala Ala Thr Leu Ala Gly Pro Ala Val Val Pro
 115 120 125
 Pro Gln Tyr Tyr Gly Val Pro Trp Gly Val Tyr Pro Ala Asn Leu Phe
 130 135 140
 Gln Gln Gln Ala Ala Ala Ala Asn Asn Thr Ala Ser Gln Gln Ala
 145 150 155 160
 Ala Ser Gln Ala Gln Pro Gly Gln Gln Gln Val Leu Arg Ala Gly Ala
 165 170 175
 Gly Gln Arg Pro Leu Thr Pro Asn Gln Gly Gln Gln Gly Gln Gln Ala
 180 185 190
 Glu Ser Leu Ala Ala Ala Ala Ala Asn Pro Thr Leu Ala Phe Gly

Ile Ser Ser Asp Gln Gln Ser Glu Met Val Lys Glu Leu Asp Gly His
 610 615 620
 Val Leu Lys Cys Val Lys Asp Gln Asn Gly Asn His Val Val Gln Lys
 625 630 635 640
 Cys Ile Glu Cys Val Gln Pro Gln Ser Leu Gln Phe Ile Ile Asp Ala
 645 650 655
 Phe Lys Gly Gln Val Phe Val Leu Ser Thr His Pro Tyr Gly Cys Arg
 660 665 670
 Val Ile Gln Arg Ile Leu Glu His Cys Thr Ala Glu Gln Thr Leu Pro
 675 680 685
 Ile Leu Glu Glu Leu His Gln His Thr Glu Gln Leu Val Gln Asp Gln
 690 695 700
 Tyr Gly Asn Tyr Val Ile Gln His Val Leu Glu His Gly Arg Pro Glu
 705 710 715 720
 Asp Lys Ser Lys Ile Val Ser Glu Ile Arg Gly Lys Val Leu Ala Leu
 725 730 735
 Ser Gln His Lys Phe Ala Ser Asn Val Val Glu Lys Cys Val Thr His
 740 745 750
 Ala Ser Arg Ala Glu Arg Ala Leu Leu Ile Asp Glu Val Cys Cys Gln
 755 760 765
 Asn Asp Gly Pro His Ser Ala Leu Tyr Thr Met Met Lys Asp Gln Tyr
 770 775 780
 Ala Asn Tyr Val Val Gln Lys Met Ile Asp Met Ala Glu Pro Ala Gln
 785 790 795 800
 Arg Lys Ile Ile Met His Lys Ile Arg Pro His Ile Thr Thr Leu Arg
 805 810 815
 Lys Tyr Thr Tyr Gly Lys His Ile Leu Ala Lys Leu Glu Lys Tyr Tyr
 820 825 830
 Leu Lys Asn Ser Pro Asp Leu Gly Pro Ile Gly Gly Pro Pro Asn Gly
 835 840 845
 Met Leu
 850

SEQ ID NO:3
 SEQUENCE LENGTH: 50
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: other nucleic acid, synthesized DNA
 SEQUENCE DESCRIPTION
 CTCTAGAGGC GGCCGCTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT

50

SEQ ID NO:4
 SEQUENCE LENGTH: 17
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: other nucleic acid, synthesized DNA
 SEQUENCE DESCRIPTION
 GTTTTTTTTT TTTTTC

17

SEQ ID NO:5
SEQUENCE LENGTH: 10
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid, synthesized DNA
SEQUENCE DESCRIPTION
AGGGCGTAAG

10

SEQ ID NO:6
SEQUENCE LENGTH: 22
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid, synthesized DNA
SEQUENCE DESCRIPTION
GCATGCCTAT TGGACTTGGG GA

22

SEQ ID NO:7
SEQUENCE LENGTH: 24
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid, synthesized DNA
SEQUENCE DESCRIPTION
GTTGGCAGTT CTGCAAGTAG TAGT

24

SEQ ID NO:8
SEQUENCE LENGTH: 22
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid, synthesized DNA
SEQUENCE DESCRIPTION
CCCATCACCA TCTTCCAGGA GC

22

SEQ ID NO:9
SEQUENCE LENGTH: 26
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid, synthesized DNA
SEQUENCE DESCRIPTION
TTCACCACCT TCTTGATGTC ATCATA

26

Subtotal